N95- 27391

5/6-82
47266
p. 12

Linear And Nonlinear Trending And Prediction For AVHRR Time Series Data

J. Smid
Department of Mathematics
Morgan State University
Baltimore, MD 21219
e-mail smid@highwire.gsfc.nasa.gov

P. Volf, M. Slama
The Czech Academy of Sciences
Pod Vodarenskou Vezi 4
Prague 8, Czech Republic
e-mail volf@utia.cas.cz
e-mail marek@uivt.cas.cz

M. Palus
Santa Fe Institute
Old Pecos Trail
Santa Fe, NM 87505
&
The Czech Academy of Sciences
e-mail mp@santafe.edu

Abstract

The variability of AVHRR calibration coefficients in time was analysed using algorithms of linear and non-linear time series analysis. Specifically we have used the spline trend modeling, autoregressive process analysis, incremental neural network learning algorithm and redundancy functional testing. The analysis performed on available AVHRR data sets revealed that (1) the calibration data have non-linear dependencies, (2) the calibration data depend strongly on the target temperature, (3) both calibration coefficients and the temperature time series can be modeled, in the first approximation, as autonomous dynamical systems, (4) the high frequency residuals of the analysed data sets can be best modeled as an autoregressive process of the 10th degree.

We have dealt with a non-linear identification problem and the problem of noise filtering (data smoothing). The system identification and filtering are significant problems for AVHRR data sets. The algorithms oulined in this study can be used for the future EOS missions. Prediction and smoothing algorithms for time series of calibration data provide a functional characterization of the data. Those algorithms can be particularly useful when calibration data are incomplete or sparse.

1. Introduction

EOS-Earth Observing System

The suite of instruments to be flown on the Earth Observing System is intended to provide a comprehensive data set of global observations of the Earth in a broad range of sensor wavelenghts. Global coverage sensors will be the primary tools for data collecting. Those instruments will play a major role in the global studies of our environment under the EOS project. To provide a quality data set for scientists, the raw data sets need to be calibrated. Obviously this effort requires a good calibration algorithm. We will focus in this paper on analysis of time series of calibration coefficients. This analysis can help in defining an optimum value of calibration coefficient based on the previous and current values determined from on-board calibration measurements. Also, the description of calibration data sets in terms of e.g. nonlinear dynamical systems and autoregressive processes will help in improving both the short and longterm estimation accuracy of calibration coefficients.

AVHRR Calibration Coefficients Data

Data sets from Advanced Very High Resolution Radiometer (AVHRR) sets were analysed. The AVHRR system was flown on the National Oceanographic and Atmospheric Administration (NOAA) operational meteorological satellites (NOAA-9/11). At a given time, the calibration data from one channel consists of a pair of numbers, the slope and intercept, that represent a straight line, the calibration curve. For one channel we obtain two time series, the slope as a function of time, and the intercept as a function of time. We have analysed the available data sets: year 1988, day 25; and year 1990, day 183.

The Calibration Time Series

The calibration data sets form time series. A time series $\{y(t)\}$ can be thought of as a realization of a stochastic process. A stochastic process can be described as a sequence of random variables. The best studied are linear stochastic processes. We used standard modeling techniques [Box, Jenkins 1976], based on autoregressive (AR) models to analyse the high frequency components of the data. Among our long term goals is to establish a short term, medium term and long term model for the estimate of calibration coefficients for AVHRR data. We would like to approximate calibration time series by linear models whenever we can avoid using computationally costly non-linear models. Testing for nonlinearity is not a trivial task. Recently tests for nonlinearity based on the redundancy (entropy) functionals have been introduced [Palus, 1993] as a tool for detecting chaotic dynamical systems. We used redundancy functionals to detect nonlinear dependencies of the time series in our data sets. Knowing that we are dealing with non-linear models we would like to build a nonlinear prediction model for calibration coefficients time series as a multivariable function of external variables (the principal one is undoubtedly the target temperature). At this phase of the analysis we have used an incremental (recursive) neural network architecture to simulate the dynamics of the behavior of calibration coefficients and the target temperature for one orbit worth of data.

We have observed four main signal components in the data sets. The first component, a slow trend, corresponds to the aging of the sensors. The second trend, a nonlinear (pseudo-periodic) dependence, corresponds mainly to the effect of the day or night part of the spacecraft orbit on the sensors (the sensors are housed in the spacecraft which is affected by variable external pseudo-periodic conditions)

The third component consists of the medium frequency signal (a few minute period).

The calibration coefficients are contaminated by noise.

2. The Time Series Modeling Goals

Data Averaging

The first goal of our analysis of the AVHRR data sets was to define the current value of a calibration coefficient, i.e. the value at a given time. Since a data set of calibration coefficients is a time series contaminated with noise, this task consists of a trend estimation, by a trend modeling and a data smoothing procedure. The results of this process are time series of calibration coefficients with noise filtered out.This process includes the estimates of confidence intervals.

Analysis of Residuals

The second goal, and a part of the data averaging procedure, is analysis of the residuals. As the series can have autocorrelated terms, it needs to be modeled by an autoregressive process.

Calibration Data Dynamics

The third goal was to investigate the modeling and predictibility of the calibration coefficients in time in terms of a nonlinear dynamical system

$$dx/dt = f(x(t), t, u(t)) \tag{1}$$

where $x(t)$ represents a time series, $f()$ is an unnknown non-linear function which we want to model and approximate, $u(t)$ represent external parameters (e.g. the target temperature).

Dynamics Diagnostic

A number of tests for non-linearity have been proposed [Tong 1993]. We have used a redundancy-based test for non-linearity [Palus 1993].

AVHRR Applications

Applications of modeling techniques for the AVHRR data sets are far reaching. Firstly, they will allow us, on a rigorous basis, to define the current calibration values of sensors more accurately and statistically fully qualified in terms of confidence intervals. Secondly, they will improve the long term and short term estimates of calibration coefficients through capturing the system dynamics.

3. Calibration Value Estimation Procedure

For AVHRR data sets, we made the assumption that the observed series can be described by the model

$$x(t) = q(t) + v(t) + e(t) \tag{2}$$

where t is the discrete reference time, $q(t)$ represents a long-term trend (pseudo-oscilations with period equal approximately to 102 minutes, or 1 orbit), $v(t)$ is a component describing the short-term periodicities. Finally, e(t) is white noise. The first term $q(t)$ is driven by external phenomena (e.g. temperature, light intensity), and the dynamics of the sensors.

A number of methods are available for signal component modeling. We focused in the present study on the method of splines, autoregressive modeling, redundancy functionals and adaptive neural network methods.

Trend Modeling in the Time Domain (Splines)

The first task, before we can analyse noise, consists of estimating and removing the component $q(t)$ from the equation (2). A splines smoothing is the standard modeling choice [Wegman 1983]. As the character of the component q(t) changes several times during one period, an algorithm is needed to capture this change in the trend. The number of knots can be controlled by the Akaike's AIC criterion [Eubank, Speckman 1990].

Dynamics Modeling in the Time Domain (Dynamical Systems)

The physical mechanisms governing the evolution of the system of calibration coefficients are not fully understood but we assumed that the data is produced by an underlying generator which is a low dimensional dynamical system.

The first component $q(t)$ is driven by external phenomena (e.g. temperature, the light intensity). This relation is complex and non-linear. In the first approximation we assumed that the *2nd order autonomous difference equations*

$$dx/dt = f(x(t)) \tag{3}$$

can simulate a dynamical system that governs the calibration coefficients. In other words we assumed that the calibration coefficients are fully described by the dynamics of the temperature evolution, which is , for simplicity, described by a *2nd order autonomous syatem*. The $2nd$ order calibration coefficient system generates a nonlinear map

$$g : S \subseteq R^2 \rightarrow R$$

$$x(i+1) = g(x(i), x(i-1))$$

where the set S contains the training exemplars. This map can be captured by a neural network algorithm.

Dynamics Modeling in the Time Domain (Neural Networks)

We have simulated one orbit worth of the 1988 data set by the Cascor learning algorithm. Because we assumed the $2nd$ order dynamics (see the previous section) we set the number of inputs for our neural net to two. Tha trained network creates a delay map characterizing the underlying dynamical system. Many different architectures of feedforward neural networks have been used for signal modeling and prediction [Chen, Billings, 1992]. Nonlinear system identification and modeling using neural networks has become very popular tool for modeling and identification of non-linear autonomous systems [Vemuri, 1994]. The neural network algorithms offer the flexibility of infinitely many parameters. The number of parameters is usually indirectly controlled through cross-validation. The method of cross-validation consists, in practical terms, of two phases. First a net is trained on a subset of data, secondly the trained net is tested on another data set. This process is repeated until the residual error on the test data hits its minimum. In our numerical simulation the cross validation test has been implemented using the Cascor code [Fahlman,1992]. The follow-on tests will demonstrate the numerical agreement of the prediction and the real data. The real data for additinal orbits are being prepared for neural network prediction tests.

Series Prediction by Bootstrapping

The trained net can be used for predicting the future values of the time series. We have used trained nets to predict various dynamic systems. Many authors have used different neural network architectures to predict chaotic systems. The standard test cases are logistic equation, Mackey-Glass nonlinear differential delayed equation and van der Pol equation.

The methodology of predicting, known as bootstrapping, works as follows: For an n input, m output network, a training exemplar is formed by taking n+m consecutive values from the data series to be extrapolated (predicted). Starting at an arbitrary value $x(i)$, the first n values $(x(i), ..., x(n+i-1))$ are presented to the network inputs. The target values are the next m values, thus a general training exemplar for an n input, m output network can be represented as $(x(i), ..., x(n+i-1)..., x(n+m+i-1))$. Each successive exemplar is formed by starting one value beyond the previous starting value. The number of exemplars sufficient to provide the desired accuracy is to be determined by numerical experiments. After the iterative learning procedure has converged we have a map (e.g. for one output, m=1)

$$x(i+1) = g(x(i), ..., x(i-n+1))$$

By bootstrapping the net into the future , this map can be iterated to give

$$x(i+2) = g(g((xi), ..., x(i-n+1)), x(i), ..., x(i-n+2))$$

The trained network can be analysed and difference equations describing the underlying dynamics can be recovered [Lowe, Webb 1994].

Noise Analysis in the Time Domain (AR-Autoregression)

The AR is the standard method for modeling linear dynamical systems. In the next step we have to decide about the character of the second component $v(t)$. The standard methodology recommends analysing spectrum and periodogram, in order to decide between the trigonometric regression or autoregression. The FFT transfomation, Fig.3, did not show well isolated high-frequency components of our signal. Therefore the standard methods using low-pass filters for noise removal are not directly applicable for our data sets. The periodogram revealed several isolated peaks in some regions. This feature indicated that trigonometric functions might be a good candidate for modeling of our data set. However, we have not observed this behavior of periodograms through the entire time domain. On the other hand, the partial autocorrelations showed significant dependencies of the lagged vector components for lags up to 10, less significant dependencies for the higher lags. That is why the AR approach has been preferred for statistical analysis of the series. In other words, if we denote the residuals of the first step of analysis $r(t) = x(t) - q(t)$, we have the model

$$r(t) = v(t) + e(t) = \sum_{i=1}^{k} b(i)r(t-i) + e(t),$$

where $b(i)$ are the parameters and k is the order of the autoregressive process AR(k).

Nonlinearity Test

The real-world data show usually some degree of nonlinearity. Our test was based on information-theoretic (redundancy) functionals. The redundancy test is based on the fact that noise, linear and nonlinear structures of a time series are represented by qualitatively different redundancy functionals.

Testing for nonlinearity of the calibration coefficients time series was performed using redundancies. Those methods were recently proposed for testing of non-linearity of dynamical systems [Palus, 1992] and are based on the general concepts of information theory.

The linear redundancy $L(X_1, ..., X_n)$ of an n-dimensional random variable with zero mean and covariance matrix C is defined as

$$L(X_1, ..., X_n) = 1/2 \sum_{i=1}^{n} \log(c_{ii}) - 1/2 \sum_{i=1}^{n} \log(\sigma_i)$$

where c_{ii} are the diagonal elements and σ_i are the eigenvalues of the $n \times n$ covariance matrix C.

The n-dimensional (non-linear) redundancy is defined as

$$R(X_1, ..., X_n) = H(X_1) + ...H(X_n) + H(X_1, ...X_n)$$

These two functionals, as the functions of the time lag τ, provide measures that differentiate linear and non-linear structures presented in the lagged versions of the component $x(t)$. The lagged version

$$(x(t), x(t+\tau), ...x(t+(n-1)\tau))$$

of $x(t)$ is a realization of the random variable $(X_1, ..., X_n)$ where n is the embbeding dimension.

4. Data Characterization

Visual inspection of the data from different time periods (1988,1990) and for different platforms revealed common features. We have searched the available data sets for features that could explain the variability of the calibration coefficients. Two features were outstanding. The long-term component $v(t)$ behaves, in the first approximation, as an autonomous dynamical system within one orbit, Fig.6.

Also PRT counts (Platinum Resistance Thermometer), located in the word 20 of the 103-word of HRPT minor frame [Kidwell. 1991] behave as an autonomous dynamical system. The target temperature is temperature of the internal target. This temperature can be calculated from the output of four PRT counts located in words 18, 19, 20 of HRPT minor frame. The conversion of PRT counts c to absolute temperature is accomplished by

$$T(K) = \sum_{j=0}^{4} a_j c^j$$

. The patch temperature (word 21) has been constant for the inspected data sets.

5. Summary of Numerical results

Our numerical results and observations are summarized in figures 1 through 17. In the following the first two digits indicate the year, and the next three digits Julian day. The slope coefficients vary significantly over one orbit. Fig.1 shows the slope coefficient series for approximately one period (88025). Fig.2 shows the PRT counts series for 88025. The FFT for the slope data (88025) did not show any significant isolated frequencies, in Fig.3. Fig.4 shows the typical relation between a slope series and PRT counts for the same time period (88025). The PRT dynamical system delay diagram (88025) shown in Fig.5 is a typical example of a map which can be learned by a neural network. The slope dynamical system delay diagram (88025) is another example of the dynamics learnable by a net (Fig.6). The PRT series (90183) shows similar features to the series (88025) even when the time difference between these series is several years (Fig.7). The same can be said about the slope values (90183), in Fig. 8. Fig. 9 is a slope vs. PRT diagram. A data smoothing procedure is illustrated in the following three figures. Fig. 10 shows a series of intercept coefficients (88025). Fig. 11 shows the smoothed intercept (88025) data. The intercept residuals (88025) showed in Fig.12 were modeled by an autoregressive process. Fig. 13 shows the best fit of the 10th order AR(10) (88025). Tests for non-linearity are captured in the remaining four figures: Fig. 14 shows slope linear redundancy statistics, dim=2 (90183). Fig. 15 shows slope linear redundancy statistics, dim=3 (90183), Fig. 16 shows slope non-linear redundancy statistics, dim=2 (9018315), Fig. 17 shows slope non-linear redundancy statistics, dim=3 (9018315). The original slope data sets were first differentiated. The results for different lags are very similar. In both cases the difference between the linear and nonliner measures is quite obvious (a factor of 10). That clearly indicates a need for a nonlinear description of the AVHRR data sets.

6. Conclusions

We have theorized, based on our numerical simulations, that the dynamics of the coefficients (within a time frame of days) is almost a periodic process with random input variables, due to such random effects as variable cloud cover.

A possible future application target of the proposed algorithms is short term, medium term and long term data prediction for MODIS. The MODIS instrument is scheduled to be launched in 1998. One of the important tasks in the processing of MODIS data will be to determine the most accurate value of the calibration coefficients and their corresponding uncertainties, in other words to model and identify time series generated by MODIS sensors in different wavebands. This can be accomplished by combining the predited values furnished by a nonlinear (neural net, nonlinear AR) or linear model. A smoothing algorithm provides calibration values without noise. A prediction model will be useful especially when we dealing with the sparse or missing data. Another potential application is the detection of sudden degradation , as opposed to gradual aging, of a sensor. That will be characterized by an abrupt change in the dynamics of the sensor. To detect this change we may compare predicted values of the time series, generated by the sensor, with the new observed values. A discrepancy between the predicted and observed values flags a faulty sensor.

7. Future Work

We are in the process of using the proposed methodology for more extensive data sets. The more extensive data sets will allow us to demonstrate the predictive power of the algorithms and to establish their error bounds. Special attention will be given to the integration algorithms for data from different calibrators (Solar Diffuser, Spectral Radiometric Calibration Assembly, Black Body, Space View), see chapter 5 in [Guenther, 1994]. Theoretical research will cover experimenting with different neural network architectures (cascade, RBF, incremental architectures). A special effort will be devoted to recovering non-linear difference equations from the trained neural network and to establishing error bounds of nonlinear predictors.

8. Acknowledgments

The GAC (The Global Area Coverage) AVHRR data were made available by M.James, Code 902, NASA/GSFC.
P.Abel, NASA/GSFC and R.Rashkin, CSC at NASA/GSFC are acknowledged for valuable comments and helpful discussions.
E.Masuoka is acknowledged for making available the computer facility of the Laboratory of Terrestrial Physics, NASA/GSFC.
J.Barker is acknowledged for making available the computer facility of the MODIS Characterization and Support Team, NASA/GSFC.
J.Smid is supported by NASA/GSFC under Grant NAG 5-2686.

References

[S.Chen and S.A.Billings,1992]*Neural networks for nonlinear dynamics system modelling and identification* , Int.J.Control, 1992, vol. 56, no. 2, 319-346.

[G.E.P.Box and G.M.Jenkins,1976]*Time Series Analysis, Forecasting and Control,* Holden-Day, San Francisco, 1976.

[R.L.Eubank and P.Speckman,1990] *Curve fitting by polynomial-trigonometric regression,* Biometrika 77, 1990, pp.1-9.

[S. E. Fahlman and C. Lebiere, 1991]*The cascade correlation learning architecture,* Research Report CMU-CS-90-100.

[B.Guenther, 1994] The MODIS Calibration Plan, Version 1.0, March 1994.

[K.Kidwell, Edt., 1991]*NOAA Polar Orbiter User's Guide*, National Oceanic and Atmospheric Administration,July 1991.

[V. Kurkova, J.Smid, 1994] *An Incremental Architecture Algorithm for Feedforward Neural Nets* , in Computer Intensive Methods in Control and Signal Processing, Preprints of the European IEEE Workshop CMP'94, Prague, 1994.

[L.Lauritson, G.Nelson, F. Porto, 1979] *Data Extraction And Calibration Of TIROS-N/NOAA Radiometers*, National Oceanic and Atmospheric Administration Technical Memorandum NESS-107,November 1979.

[D.Lowe, A.R.Webb, 1990] *Time series prediction by adaptive networks: a dynamical systems perspective,* in V. Rao. Vemuri, R.D.Rogers,eds., 1994, *Artificial Neural Networks ,Forcasting Time Series,* IEEE Computer Society Press, 1994.

[M. Paluš, 1993] *Identifying and quantifying chaos by using information-theoretic functionals,*, in Time Series Prediction: Forecasting the Future and Understanding the Past, A.S. Weigend and N.A. Gershenfeld, (eds.), Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XV (Addison–Wesley, Reading, Mass., 1993) pp. 387-413.

[C.R.Rao, 1965]*Linear Statistical Inference and Its Applications,* Wiley, New York, 1965.

[V. Rao, Vemuri, R.D.Rogers,eds., 1994] *Artificial Neural Networks ,Forecasting Time Series,* IEEE Computer Society Press, 1994.

[H.Tong, 1993] *Non-linear Time Series, Dynamical System Approach* Clarendon Press Oxford, 1993.

[P.Volf, 1994] *Dimension Reducing Approach in Statistical Regression Analysis,* in Computer Intensive Methods in Control and Signal Processing, Preprints of the European IEEE Workshop CMP'94, Prague, 1994.

[E.J.Wegman and I.W.Wright, 1983] *Splines in statistics,* Journal of Amer. Statist. Assoc. 78, 1983, pp.351-365.

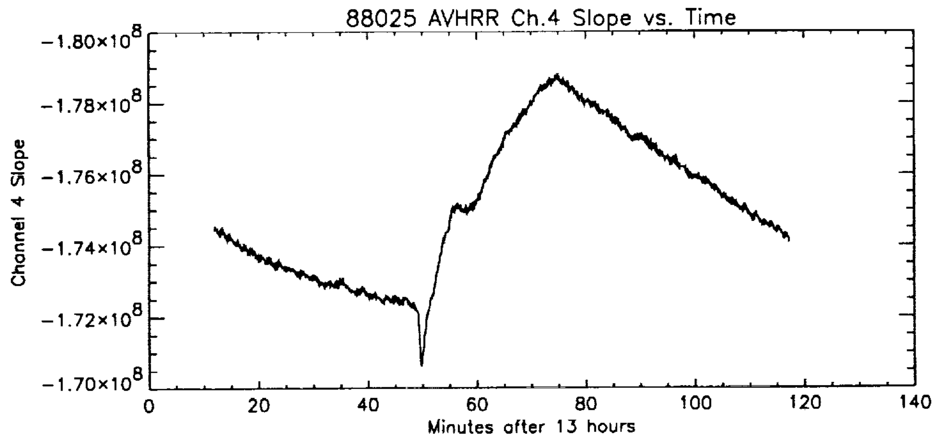

Fig.1

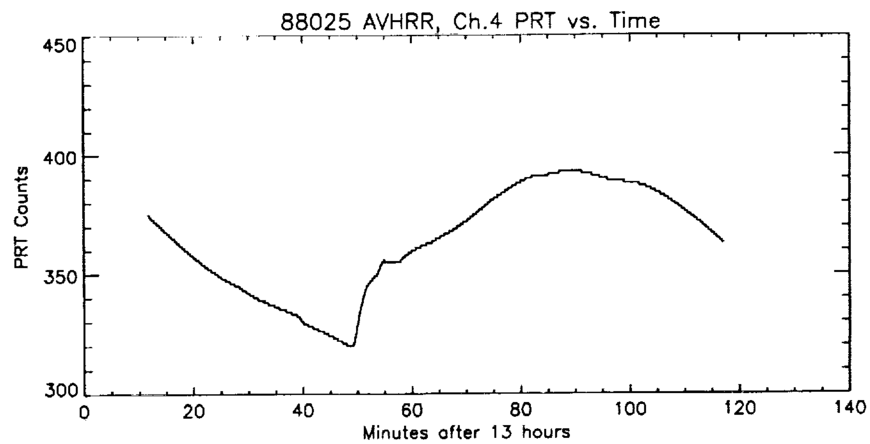

Fig.2

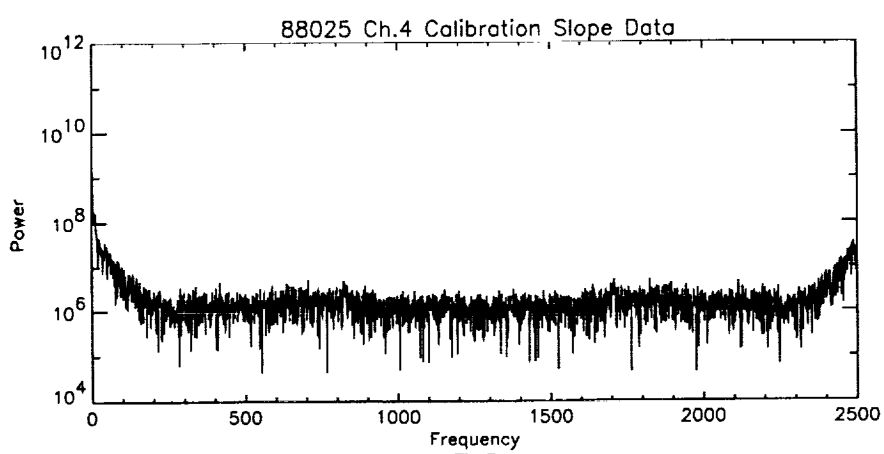

Fig.3

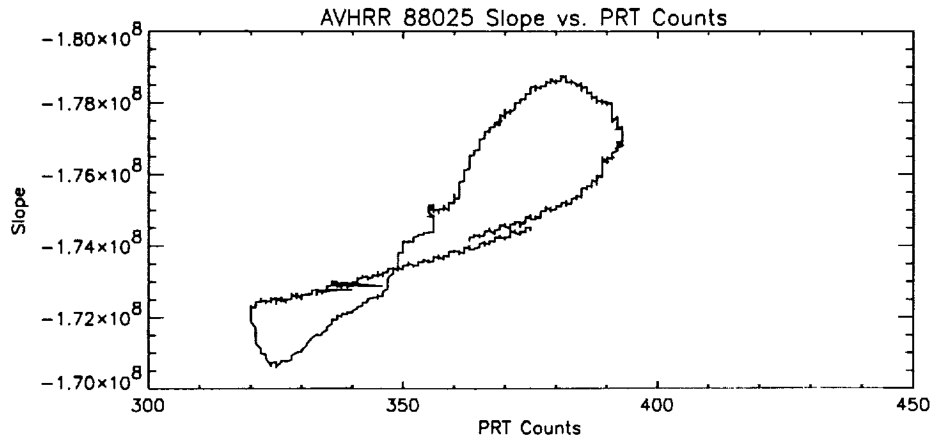

Fig.4

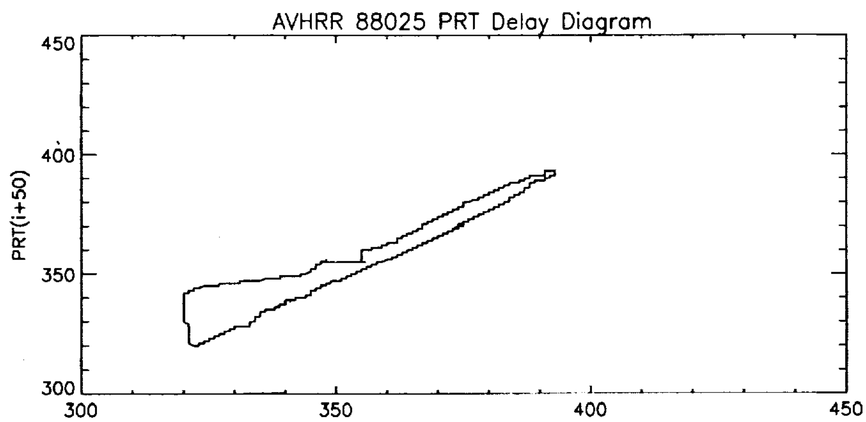

Fig.5

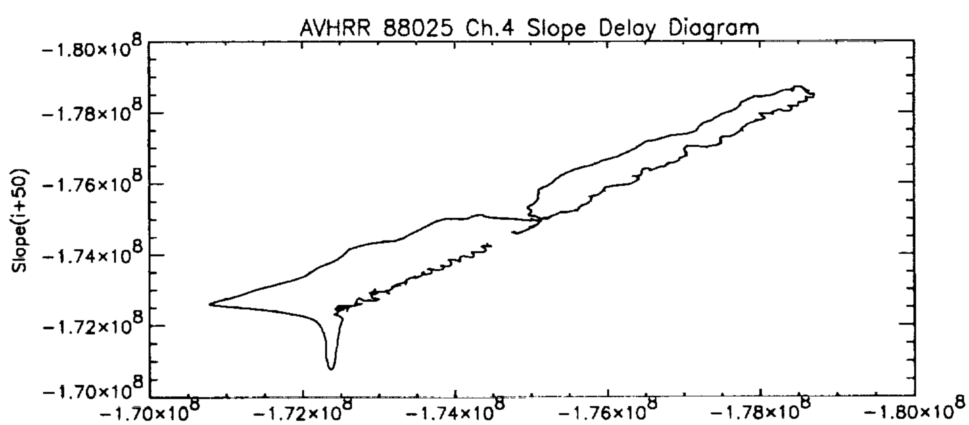

Fig.6

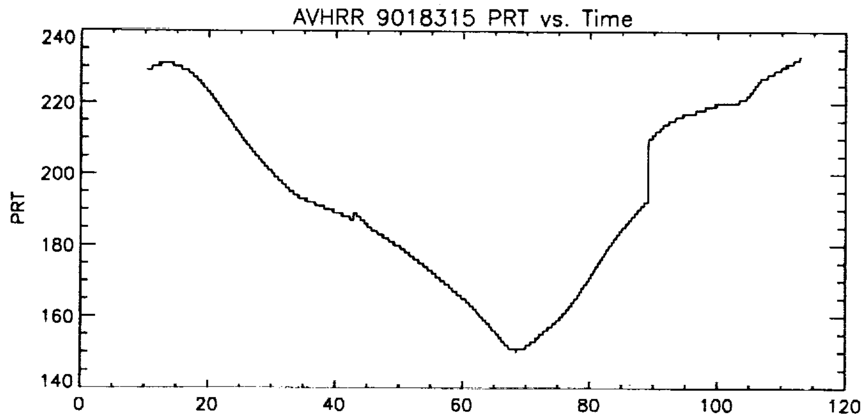

Fig.7

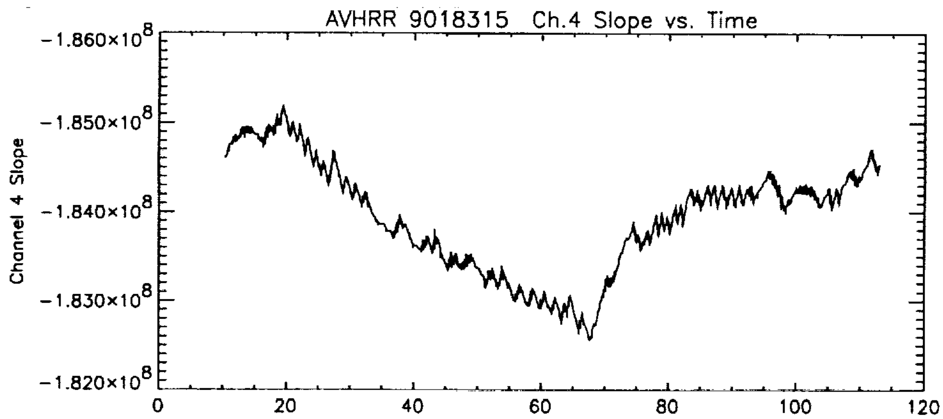

Fig.8

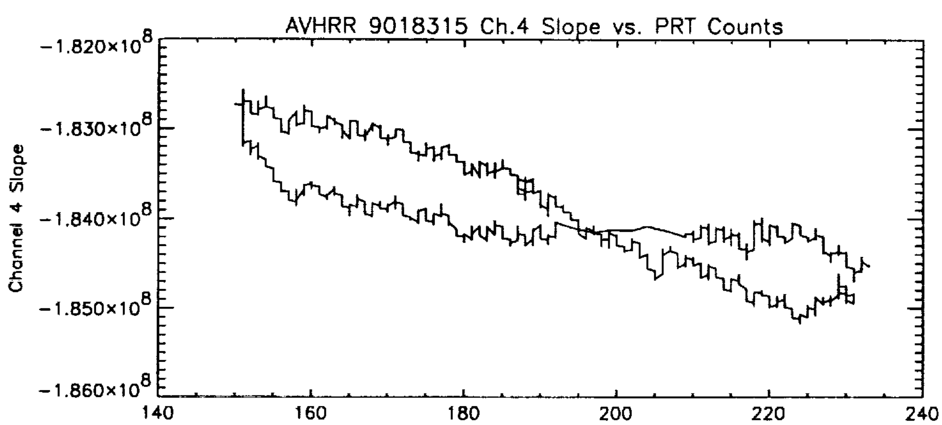

Fig.9

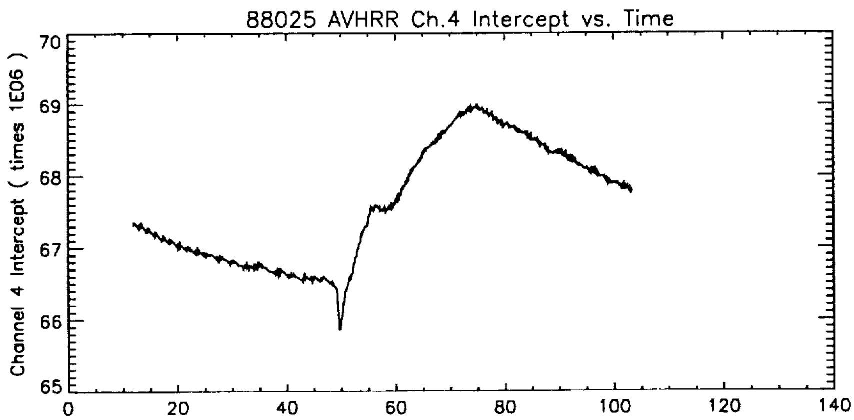

Fig.10

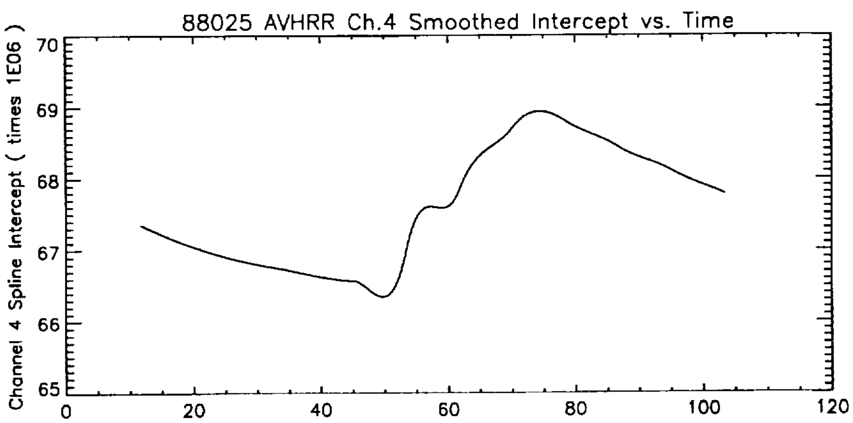

Fig.11

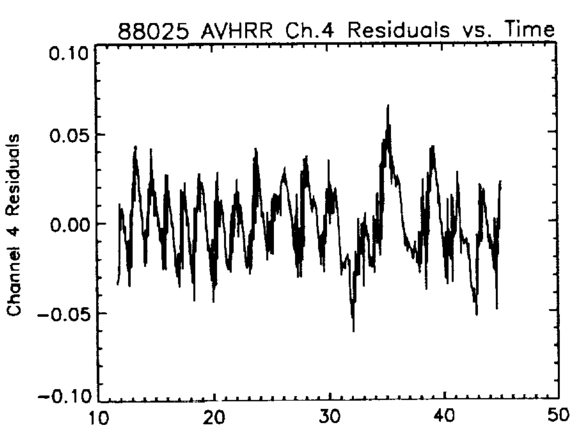

Fig.12

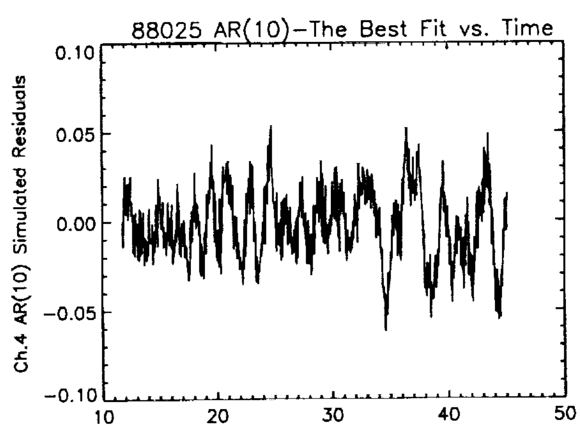

Fig.13

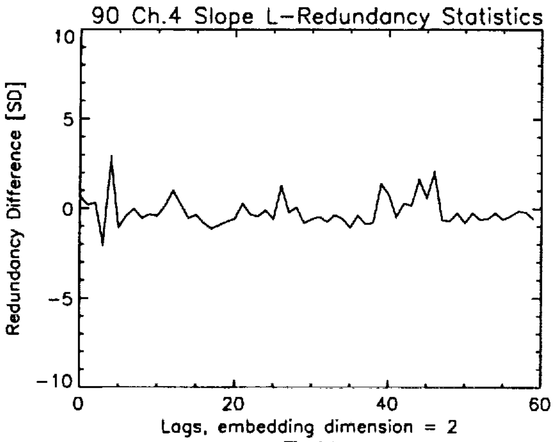

Fig.14

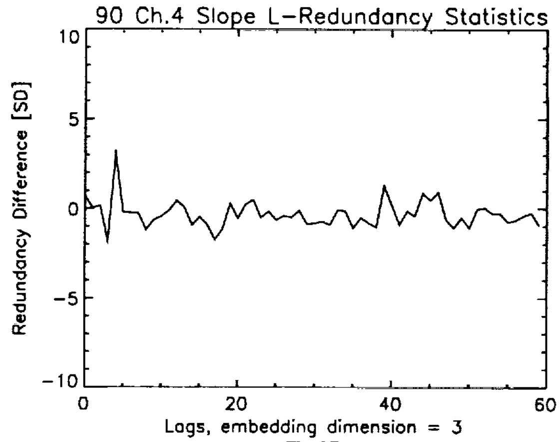

Fig.15

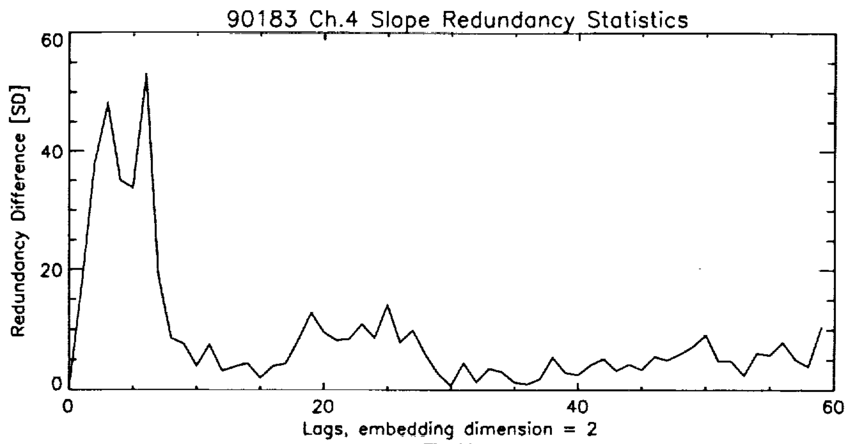

Fig.16

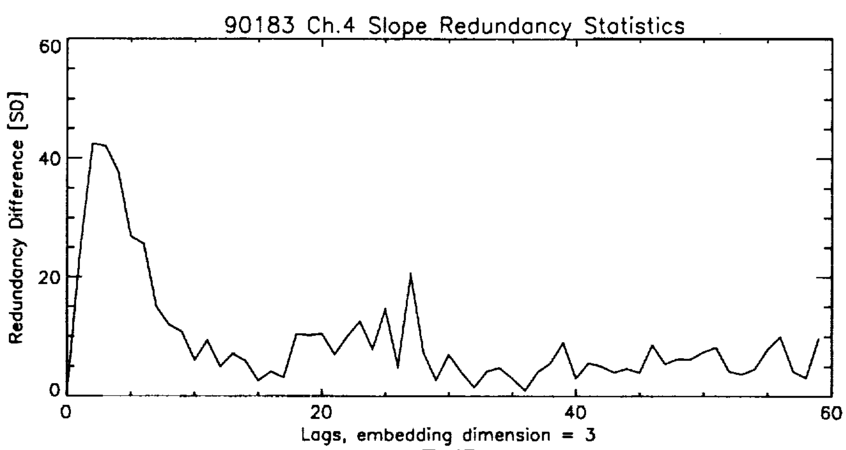

Fig.17